A neural network for electromagnetic shower energy measurement

Y. Caffari

A Neural Network for electromagnetic shower energy measurement

Y. Caffari, IPN Lyon, France

4th May 2006

Abstract

This note describes a new method based on a neural network to measure the energy of electromagnetic cascades. After applying a specific algorithm of reconstruction [1], we extract informations from the longitunal and transversal profiles to evaluate the energy of the primary electron. This neural network is tested on Monte Carlo simulations and experimental data (6 GeV, dry scan). Preliminary results are also presented.

Contents

1 Introduction

The energy is an important quantity to measure for the study of the $\nu_\mu \rightarrow \nu_\tau$ oscillations, with τ decaying in electron, and the $\nu_\mu \rightarrow \nu_e$ oscillations. This note proposed to expose a method based on a neural network for energy evaluation of electromagnetic showers. In the section 2, we describe the processing of the data and simulations. The third section explains in details the treatment of the background of the test beam. In the section 4, we remind briefly the algorithm of reconstruction used for the analysis. In the fifth section, we present the first input variable of the neural network : the number of basetracks. The two next sections are dedicated to the analysis of the longitudinal and transversal profiles. The section 8 is devoted to the results for the energy estimation by the neural network.

2 Treatment of the data and Monte Carlo simulations

2.1 Experimental data analysis

2.1.1 Experimental set up

An electron test beam has been performed at DESY in Germany using the T24 electron beam line. A detailed description of the experimental set up is given in the reference [2]. In this test, it is possible to have a very pur electron beam and the chosen density (1 e-/cm^2 or 100 e-/cm^2). In this note we decide to expose some results of a brick made of 20 emulsion layers interspaced with 1 mm thick lead plates and exposed at a high density of 6 GeV electrons.

2.1.2 Data reconstruction

Experimental data had been scanned at Neuchatel with dry objective. The microtrack reconstruction is performed on line by the European scanning system (hardware) and Sysal (Software). Microtracks with at least 6 grains and $|\tan(\theta_{xz})|$ and $|\tan(\theta_{yz})| < 1$ rad are kept. For the basetrack reconstruction off-line, FEDRA [3] is used. The basetrack efficiency is about 90 %. In this study we keep basetracks which satisfy the following cut :

- $|\tan(\theta_{xz})| < 400$ mrad and $|\tan(\theta_{yz})| < 400$ mrad, where z axis direction is defined as the perpendicular direction on the transverse plan of the brick.
- $\chi^2 < 0.333 \times n_{grain}$ - 4.343.

The tracking is realized by the algorithm described briefly in the next section.

In this test beam there are 2 important sources of errors which must be treated with special attention.

1. An important background was integrated during the test beam. First there was an important exposure to the cosmics and in a second time the time storage before and after the exposure was long (2 months in total). As a consequence the background is much higher to the one expected in the OPERA experiment.
2. The high density exposure (100 e-/cm^2) doesn't constitue an ideal case for the energy measurement. Indeed a lot of showers overlapped each other. But on our study we select isolated cascades and we limit our study inside a cylinder with a radius of 400 μm and 20 emulsion layers length. This configuration allows to test the limits of the neural network.

film number	1	2	3	4	5	6	7	8	9	10
btk density (mm^{-2})	88.7	67.4	80.4	54.8	58.0	62.2	83.2	68.7	47.9	76.7
film number	11	12	13	14	15	16	17	18	19	20
btk density (mm^{-2})	73.5	62.0	71.8	59.8	49.6	60.1	64.8	71.2	67.3	57.4

Table 1: Mean density of basetracks per film of an non-exposed area (background)

2.2 MC simulation analysis

The official OPERA framework [4] is used for the MC simulations. OpGeom package [5] allows to reproduce the modular geometry of a brick and the simulation of an electromagnetic shower is done by the GEANT 3 VMC (OpSim Package). The microtracks are reconstructed from the couples of hits produced at the entrance and at the end of an emulsion layer. The hits on the plastic base are smeared in x,y position and in θ_{xz} and θ_{yz} angle. The basetracks which obey to the following cuts are hold :

- $|\tan(\theta_{xz})| < 400$ mrad and $|\tan(\theta_{yz})| < 400$ mrad.
- The absolue value of the angle difference between microtracks and basetracks must be smaller than 100 mrad.

The specific algorithm reconstucts the showers of the simulations.

Contrary to the algorithm of reconstruction, the MC simulations are performed without background. In the rest of the note we will present first the results for simulations of electrons without background in a full brick (57 films). Then other simulations with electrons in brick made of 20 emulsions and a shower contained in a cylinder with a radius of 400 μm are useful to see the evolution of the distributions and understand their effects on the energy evaluation and resolution.

3 Treatment of the background

3.1 Background estimation

As mentioned before, an important background was integrated during the test beam. Nevertheless, a good understanding of it can reduce his impact on the systematic errors. The mean density is calculated from a 2 cm^2 non-exposed area and it is estimated at about 65 basetracks/mm^2/film. The figure 1 shows the mean background per film of the non-exposed area and it is summarized in the table 1.

3.2 Background substraction

The main idea is to be background-independant for the energy estimation. Indeed, the background introduced during the reconstruction depends on the choice of the parameters, the scanning strategy,...The systematic errors are difficult to simulate. Therefore, to train the neural network files containing signal without background are produced and can be used by everyone.

Nevertheless, before using the neural network to estimate the energy a work of background substraction must be done.

When the algorithm of reconstruction is applied, it includes fake basetracks (background) with the signal (electromagnetic shower). As a consequence the number of basetracks x_i per bin i can be written as : $x_i = s_i + b_i$, with s_i being the number of "signal basetracks" per bin i and b_i the number of "fake basetracks" (background) per bin i. The main work is to estimate b_i and to substract it correctly in order to have finally $s_i = x_i$ - b_i like in the files for training the neural network. This work is divided in 2 steps.

1. First, simulations with background is performed. So the mean number b_i of fake basetracks per bin i is known. The mean background is represented in green in the figures 8 and 18.

2. The second step consists in substracting event by event the background. But we don't use directly the mean value b_i. Indeed, the stastic fluctuations event by event on the value b_i must be take into account and we introduce the following parametrization : b_i is calculated from a gaussian distribution which mean value is equal to the mean background in a bin i.

4 Algorithm of reconstruction

The energy measurement is performed thanks to the algorithm of reconstruction presented in the note [1]. This algorithm is divided in two main steps :

1. the first step consists in reconstructing the primary track of the electron which generates the shower. The algorithm starts from the first film by considering each basetrack as a starting point for the primary track and propagates itself until the fifth film in order to connect the basetracks. During its propagation it can take into account the presence of one hole. Finally among all the possibilities and combinations, it holds one track which it considers as the best one to be the "primary track".

2. From the first basetrack of the primary track, the algorithm opens a cylinder and links all the basetracks film by film. This second step allows to build the branches of the shower.

The connection between two consecutive basetracks follows specific criteria which take into account the physic of the electromagnetic cascade (see reference [1]).

5 Mean number of basetracks : nbtk

When a high energy electron passes through lead plates, it initiates an electromagnetic shower by pair creations and bremsstrhalung ([6] , [7]). The secondary particles produce tracks in the emulsions. And the total length of the track is directly proportionnal to the energy of the primary electron. Thus, the first idea is to count the number of basetracks in the shower. The figure 2 represents the mean number of basetracks in a shower developping in a full brick. The mathematical relation linking both quantities is :

$$E_{rec} = \frac{nbtk + 5.32}{31.85}$$

The figure 4 shows the distributions of the number of basetracks for different values of the energy. If we consider development of a shower in an entire brick (57 films), the separation of the distributions is good enough and the variable “number of basetracks” nbtk is a natural candidate for the neural network. Nevertheless, the figure 5 which presents results for shower developped in 20 films and cointained in the cylinder shows a degradation in the separability of the distributions. Moreover, the figure 3 diplays the mean number of basetracks in a shower as a function of the electron energy. For high energy values, ie more than 6 GeV, we can see a deviation from the straight line. This phenomena is due to containment effects : the volume of study is not sufficient to contain all the basetracks of the shower.

Then the comparison between MC simulations and data in the figure 6 confirms the precedent results.

6 Longitudinal profile

6.1 Mean profile reconstructed

The longitudinal profile of an electromagnetic shower gives information about the energy deposition film by film and varies for different values of the energy of the primary electron. The figure 7 which presents longitudinal development for low, medium and high energy electrons illustrates this idea. The mean number of basetracks film by film increases and the position of the maximum of the distribution evolves with the energy. The mean longitudinal profile is well modelized with a gamma function :

$$\rho(n_i) = \alpha b \frac{(bn_i)^{a-1} e^{-bn_i}}{\Gamma(a)}$$

Where n_i is the emulsion number. The information on the energy can be extracted from the coefficient α and nmax (see the two next paragraphes).

The figure 8 shows a comparison between data after background substraction and MC simulations. Both experimental and theoric profiles have the same evolution. For the data we denote some fluctuations due to the low statistic (172 events).

6.2 Normalization of the longitudinal distribution : α

The coefficient α is a factor of normalization and is directly proportionnal to the energy of the primary electron as it is shown in the figure 9. The linear relation between α and the electron energy is :

$$\alpha = 31.38 \times E - 0.52$$

by considering MC simulations in a full brick. To extract α from each individual event we decide to calculate it from the mean value and the RMS of the longitudinal profile. This kind of calculation is applied for each variable described in the rest of this note. Thus, the calculation allows to take into account statistic fluctuations for each individual event. The distributions of the coefficient α for different values of the energy are displayed in figure 10. It shows that the distributions are well separated. Nonetheless, this separability is damaged when the volume of study is limited (n=20 films and r=400 μm) as it is demonstrated in the figure 11. The figure 12 shows a comparison data/MC distributions.

As a conclusion the coefficient α can be considered as an input variable for the neural network.

6.3 position of the maximum : nmax

The figure 7 demonstrates that the position of the maximum of the longitudinal profile increases with the energy value. The coefficient nmax is linked to a and b by the following relationship :

$$nmax = \frac{a-1}{b}$$

.

The figure 14 represents the distributions of the nmax coefficient for showers contained in a full brick. The distributions are seperated but this separability decreases for high values of the energy because nmax is proportional to the logarithm of the energy (see figure 13 where nmax = 3.17 × ln(E) + 10.9). The overlapping becomes more important when the study is done on electromagnetic cascades cointained inside a cylinder with a length equivalent to 20 films and with a radius of 400 μm (see figure 15). Then the figure 16 shows a good agreement between data and MC simulations.

7 Transversal profile

7.1 Mean profile reconstructed

The transversal profile describes the lateral deposit energy in an electromagnetic shower. The figure 17 shows lateral profiles for low, medium and high values of the energy and they are well modelized by the following function :

$$f(r) = C1e^{-a1r}$$

The profiles change with the value of the energy because the density of basetracks produced is correlated to the energy of the primary electron. The best way to extract and quantify this information is to check the evolution of the distributions of the coefficients C1 and a1.

The figure 18 shows a comparison between data and MC carlo simulations. We see that the first bin is underestimated for the data.

7.2 Normalization of the lateral profile : C1

The coefficient C1 is a factor of normalization and its value depends on the energy and is directly correlated to the basetrack density. This idea is illustrated in the figure 19 where :

$$C1 = 30.2 \times E - 8.97$$

The distributions of the normalization factor is illustrated in the figure 20. The separability of the distributions appears clearly and C1 can be considered as an input variable for the neural network. As for the variables nbtk, α and nmax, this separibility decreases with the volume of study. This effect is summarized in the plots of figure 21. The comparison experimental data and MC simulations appears in the figure 22.

7.3 Slope of the distribution : a1

The coefficient a1 represents the slope of the lateral profile. Its distributions for different values of the energy are shown in the figure 23 for electromagnetic showers full contained in a brick and for cascades in 20 emulsion layers and inside a cylinder (r = 400 μm) (see figure 24). The width of the distributions decreases with the energy and a1 seems to be an intersting candidate for the neural network.

The figure 25 diplays a comparison between MC simulations without background and experimental data with coefficient calculations after backgound substraction.

8 Energy estimation by the neural network

8.1 Description of the neural network

In the proposal, a method is explained to evaluate the energy of an electromagnetic cascade ([9]). By counting the number of basetracks inside a cone with an angle aperture of 50 mrad, it's possible to reconstruct the momentum (or energy) of the primary electron. But another method presented in this note proposes to use a specific neural network (NN).

The energy estimation is based on a Multilayer Perceptron Neural Network developped under the ROOT framework . A detailed description is made in the ROOT documentation (see reference [8]). The NN is divided in 3 layers :

- The first layer is composed of 5 neurons. Each neuron is linked to an input variable, i.e. nbtk, α, nmax, C1 and a1.
- the second layer, the hidden layer, is made of 5 neurons too. Each of these neurons is connected to other one of the precedent layer (so 5 $\times$ 5 connections).
- the last layer has one neuron and is called the output layer. Its role is to give the value of the energy from the value of the 5 coefficients. The output neuron is connected to the five ones of the hidden layers.

10000 events are simulated in a brick without background in a range from 1 GeV to 10 GeV with 1 GeV steps (1000 events by energy value). From these files, input variables are calculated and used to train the NN.

8.2 Energy evaluation and resolution

The NN was first tested with GEANT3 Monte Carlo simulations of electromagnetic showers contained in a full brick. The results are visible in the figure 26, figure 27 and figure 28. The energy is well reconstructed by the NN and with a resolution :

$$\frac{\Delta E}{E} = \frac{24.6(\%)}{\sqrt{E(GeV)}}$$

Then the figure 29 shows a comparison between the energy reconstructed as a function of the number of plates and the true value of the energy (6 GeV). The energy estimation gets closer to 6 GeV (true value) when the number of plates increases. Moreover, the resolution becomes better with the number of plates. This can be seen on the figure 30 where the σ of the gaussian decreases with the number of plates.

The figure 31 compares energy estimation from data after background substraction and MC simulations without background. We recall both experimental data and simulations are studied in limited conditions. The values are summarized in the table 2.

	data with background substraction	MC without background
Energy evaluation (in GeV)	6.12	6.32
RMS of the distribution	1.90	1.39
Energy resolution	31 %	22 %

Table 2: Energy estimation and energy resolution for data with background substraction and MC simulations without background. The study is realized in 20 emulsions and inside a cylinder with a radius of 400 μm.

9 Conclusion and perspectives

In this note, we have presented a method based on a neural network to measure the energy of an electromagnetic shower. The neural network uses 5 input variables extracted for the calulation of coefficients based on the number of basetracks, the longitudinal and transversal profiles. The results presented in this note are preliminary. They need to be completed with other values of the energy and a test beam with conditions more closer than the OPERA experiment. Thus, a better agreement between MC simulations and experimental data can be reached. Futhermore another complementary algortihm can be used to test the neural network. Other studies for the electron/pion separation are based on the same neural network.

References

[1] Y.Caffari, F.Juget, Algorithm of reconstruction for electromagnetic shower analysis in emulsion cloud chambers, 2006, OPERA internal note.

[2] D.Autiero, Y.Caffari, L.S. Esposito, A. Marotta, J. Marteau and P. Migliozzi, Characterization of T24 the electron beam line available at DESY, April 7 2004, OPERA internal note.

[3] http://nts-lab01.na.infn.it/fedra/

[4] http://operaweb.web.cern.ch/OperaWeb/internal/exchanger/home/software/ documentation/OpRelease/ (internal section of the opera home page)

[5] L. Chaussard, OpGeom, software model of the opera detector, July 1 2005, OPERA internal note.

[6] Particle data group, Particle Physics Booklet, July 2004.

[7] B. Slowinski, Phys. Part. Nucl. 25 (2), March-April 1994

[8] René Brun and al., ROOT documentation version 4.02, 2005.

[9] The OPERA Collaboration, OPERA Proposal, CERN/SPS 2000-28, SPSC/P318, LNGS P25/2000.

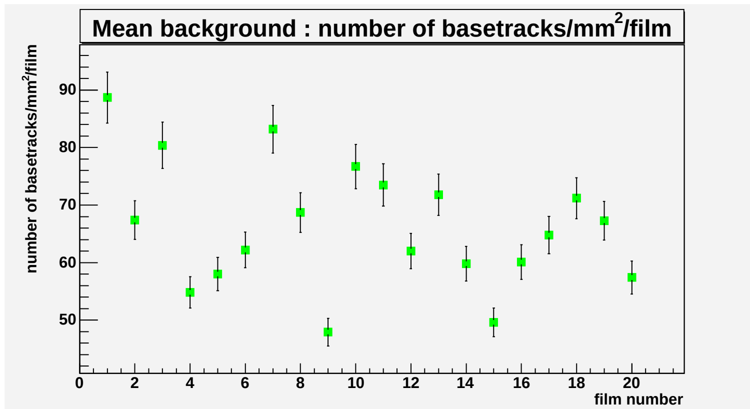

Figure 1: Background in a non-exposed area (2 cm^2)

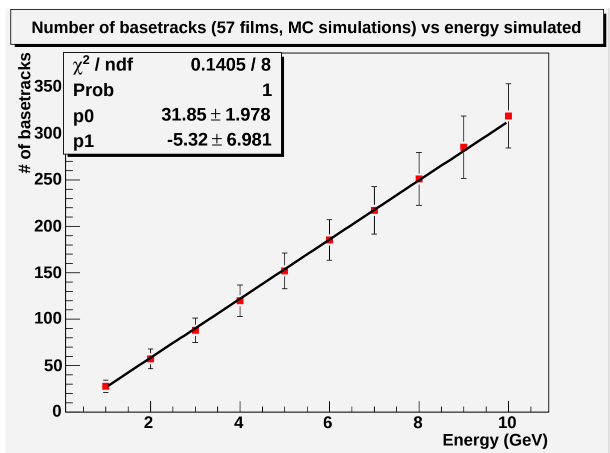

Figure 2: Mean number of basetracks as a function of the energy of the electron (MC simulations without background, 57 films).

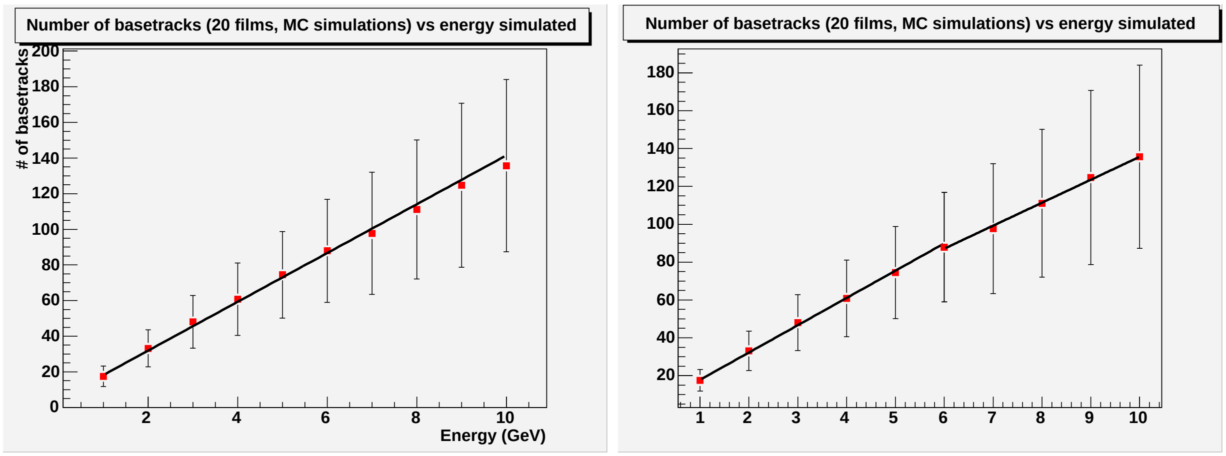

Figure 3: Mean number of basetracks as a function of the electron energy (MC simulations without background, 20 films). The left plot shows a slight deviation from the straight line for high energy points (from 6 GeV to 10 GeV). The right plot shows that 2 linear relations link the mean number of basetracks and the electron energy (see the text).

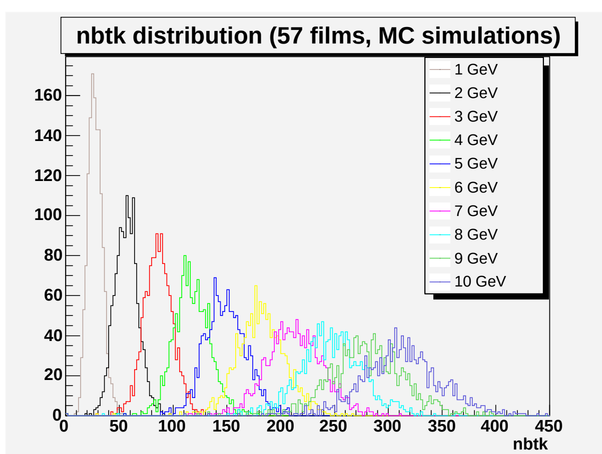

Figure 4: Distributions of the number of basetracks (simulations without background, 57 films)

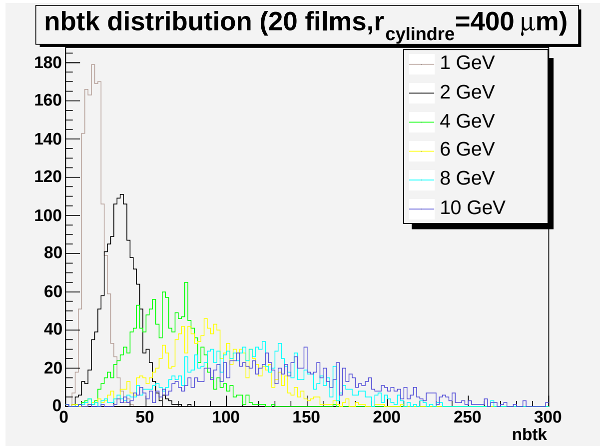

Figure 5: Distributions of the number of basetracks (20 films, r=400 μm, 6 GeV).

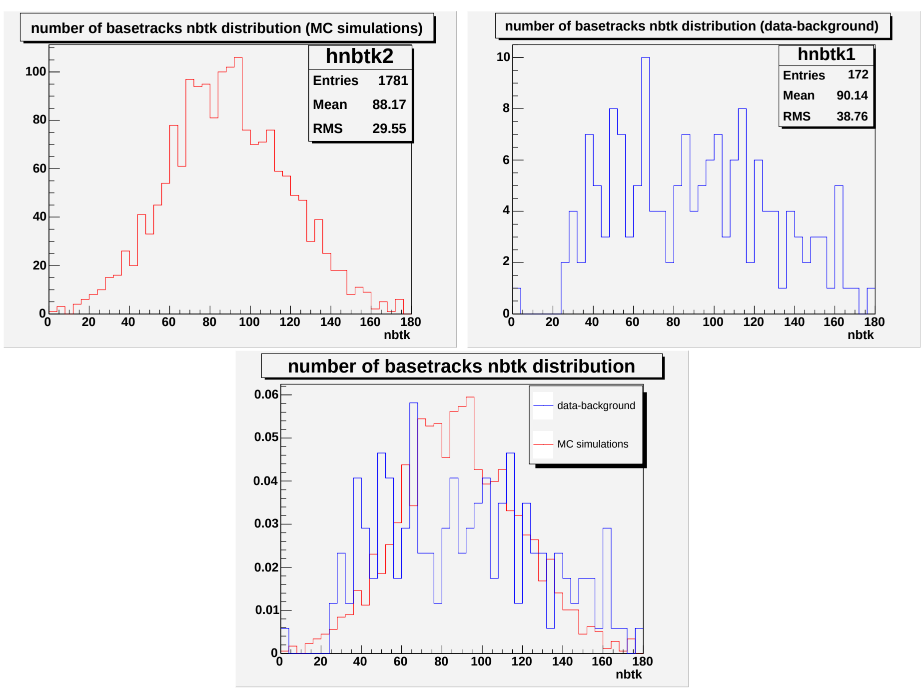

Figure 6: Distributions of the number of basetracks (data after background substraction, 20 films, r=400 μm)

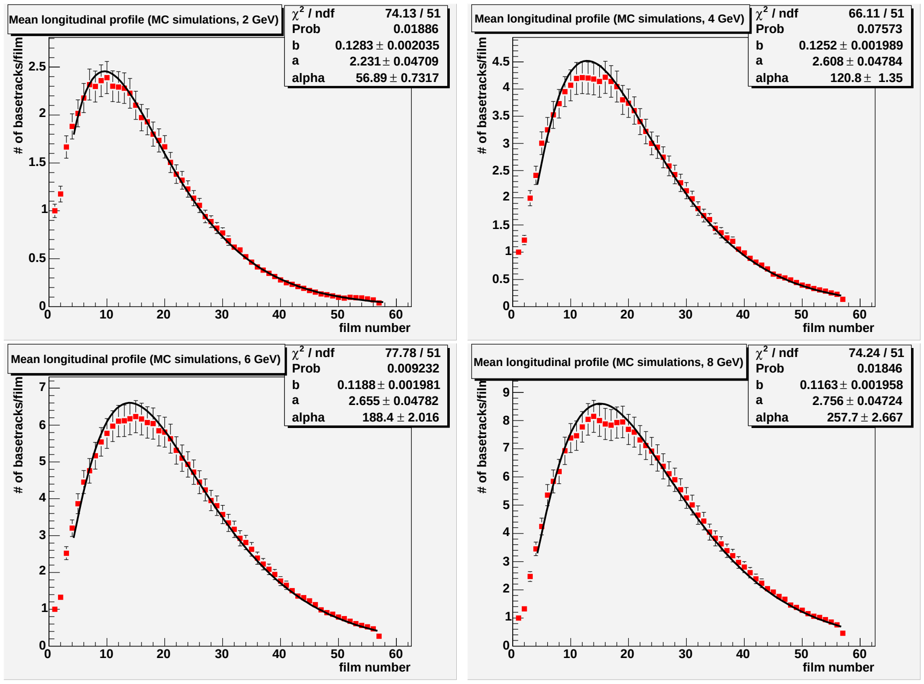

Figure 7: Mean longitudinal profile (57 films, MC simulations without background)

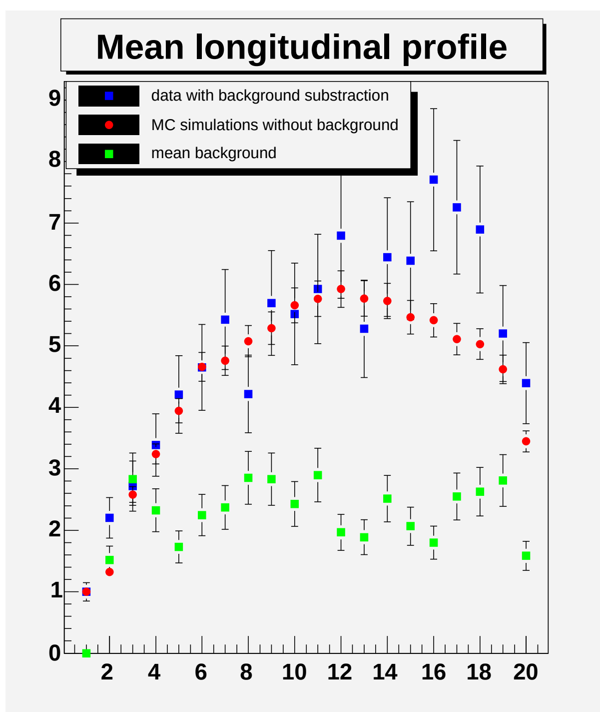

Figure 8: Mean longitudinal profile of the electromagnetic cascade. The x-axis represents the emulsion number and y-axis the mean number of basetracks per film.

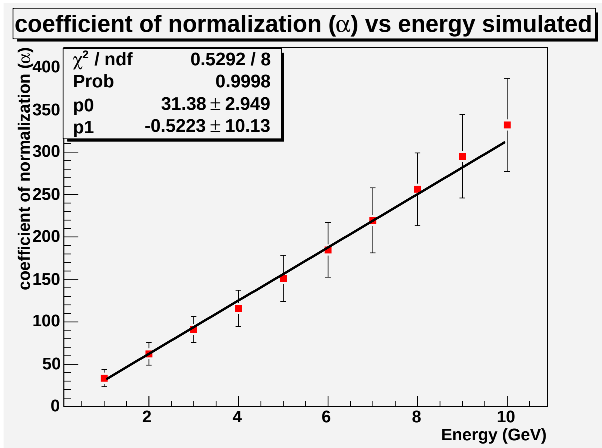

Figure 9: Mean value of the coefficient α as a function of the electron energy

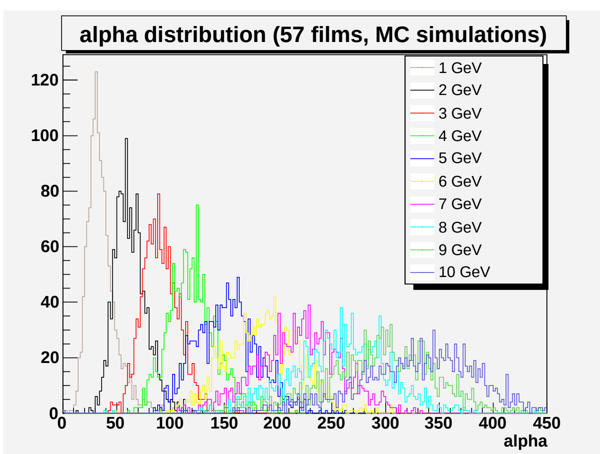

Figure 10: Distributions of the coefficient α (simulations without background, 57 films)

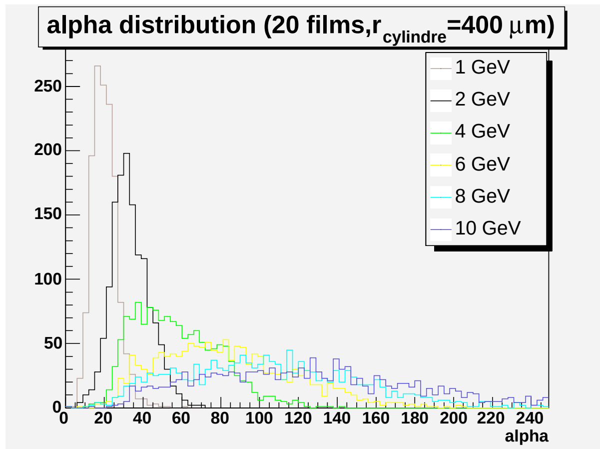

Figure 11: Distributions of the coefficient α (simulations without background, 20 films, r=400 μm)

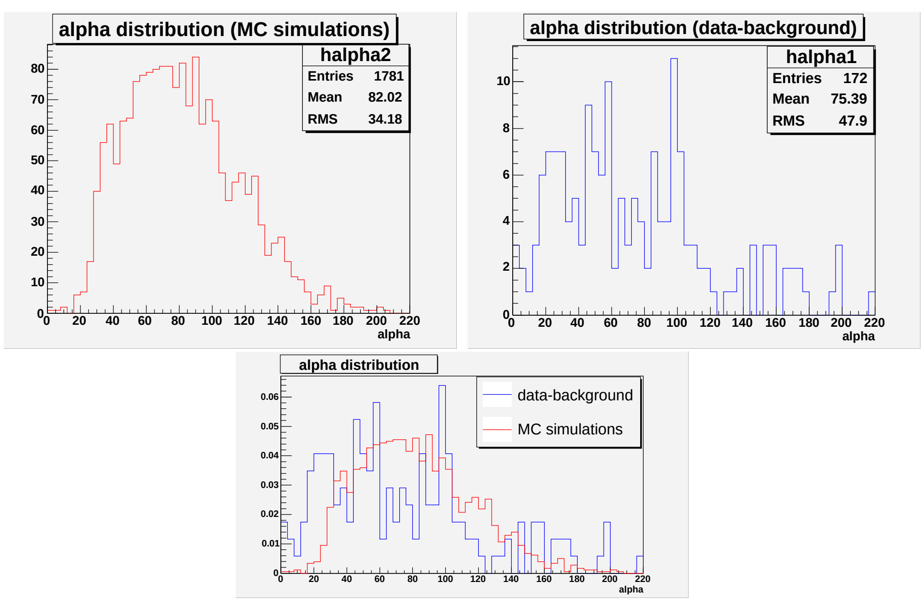

Figure 12: Distributions of the coefficient α (20 films, r=400 μm, 6 GeV). The red histogram represents the distribution of simulations withtout background and the blue histogram concerns data after background substraction.

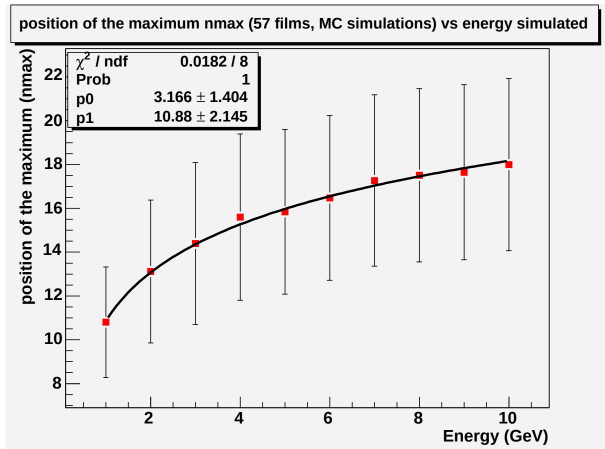

Figure 13: Mean value of nmax as a function of the electron energy (57 films, MC simulations without background)

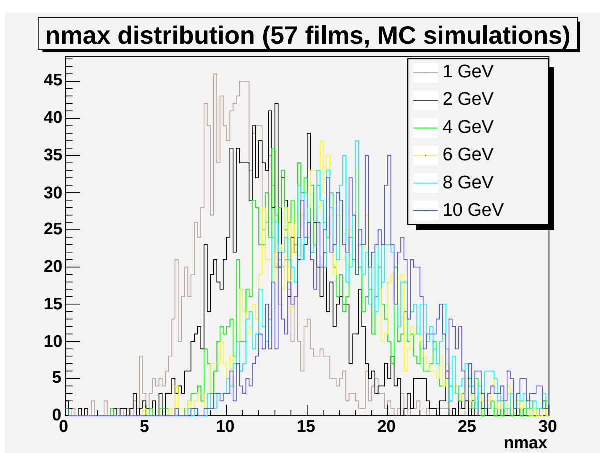

Figure 14: Distributions of the maximum position (simulations without background, 57 films)

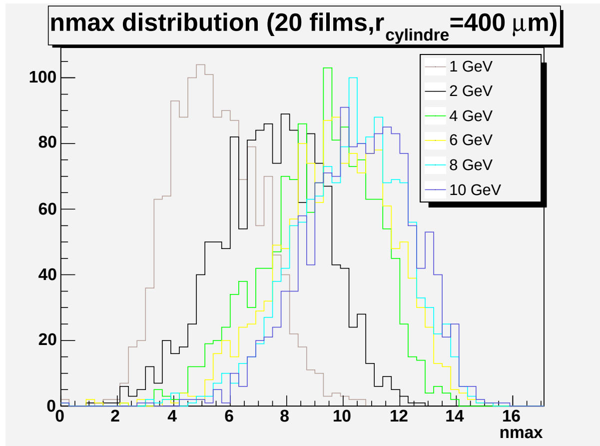

Figure 15: Distributions of the maximun position nmax (simulations without background, 20 films, r=400 μm)

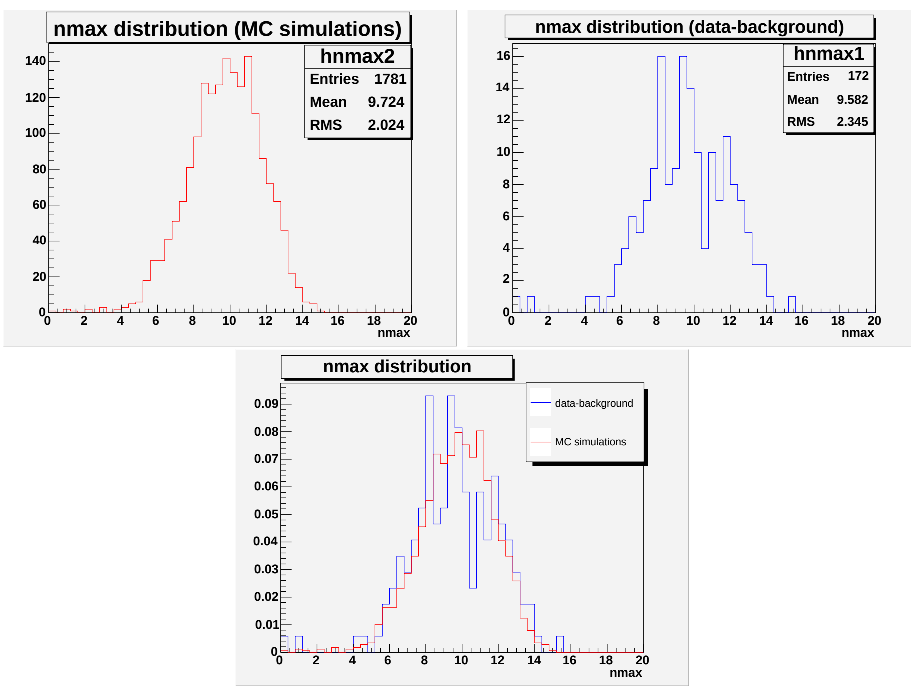

Figure 16: Distributions of the maximum position nmax (20 films, r=400 μm, 6 GeV). The red histogram represents the distribution of simulations withtout background and the blue histogram concerns data after background substraction.

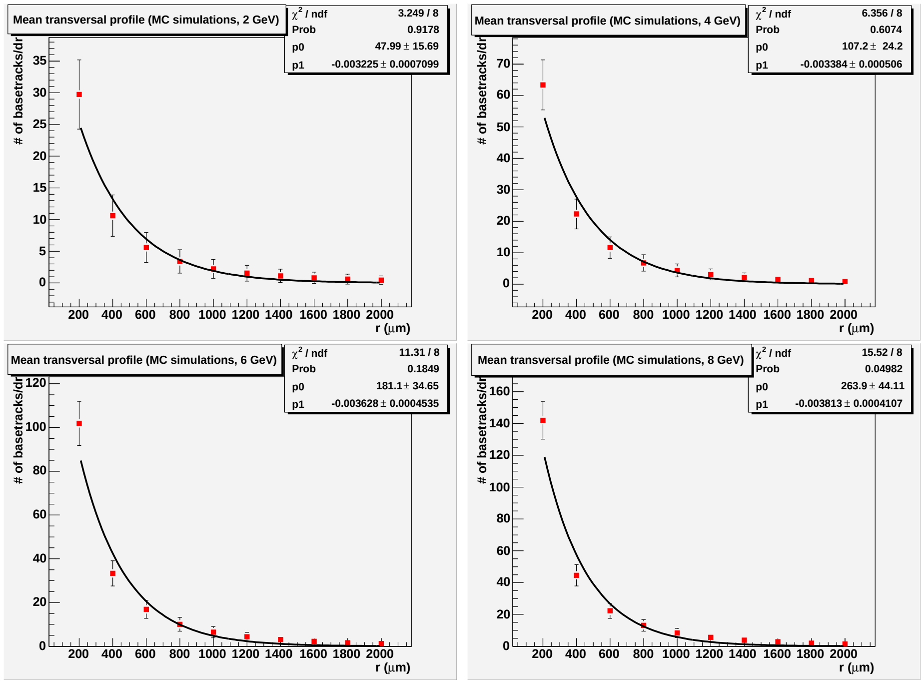

Figure 17: Mean transversal profile (57 films, MC simulations without background)

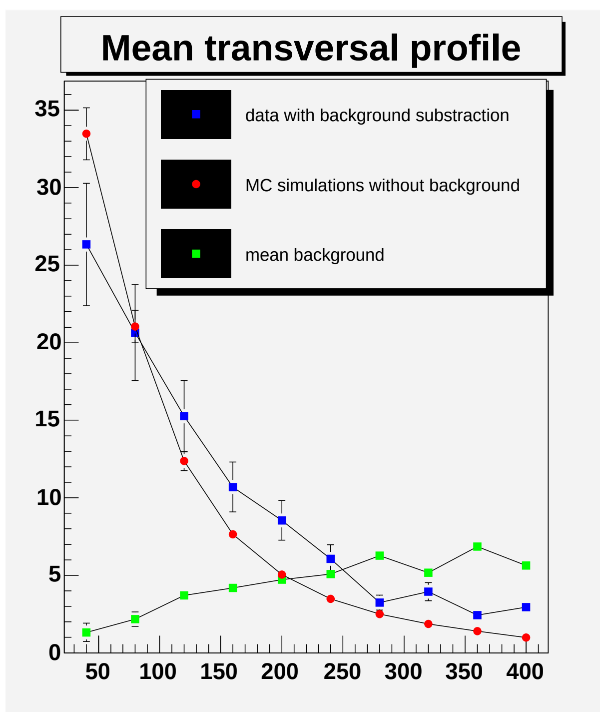

Figure 18: Mean transversal profile of the electromagnetic shower, each bin for 40 μm. The x-axis represents the radius r (μm). The y-axis represents the mean number of basetracks per bin.

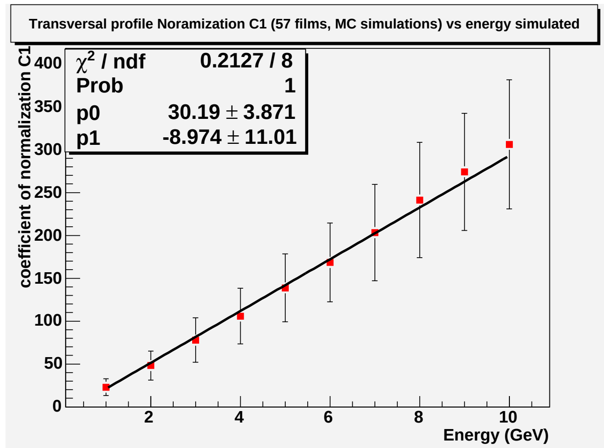

Figure 19: Mean value of the coefficient C1 as a function of the electron energy (simulations without background, 57 films)

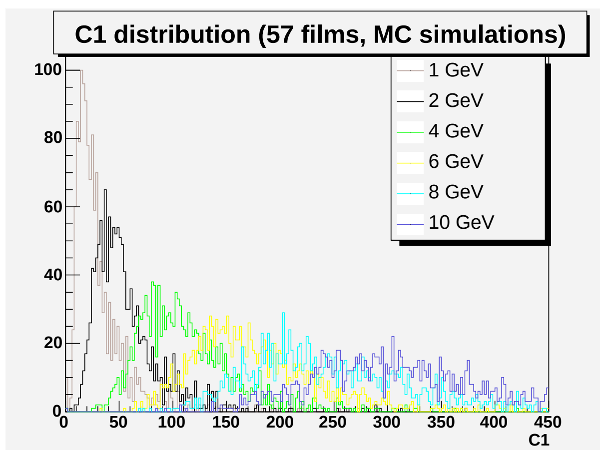

Figure 20: Distributions of the coefficient C1 (simulations without background, 57 films)

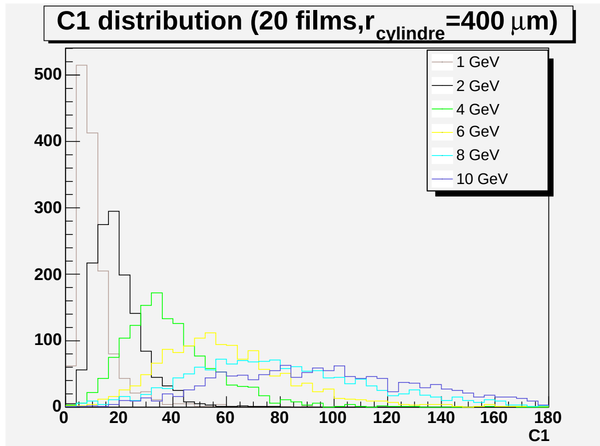

Figure 21: Distribution of the coefficient C1 (simulations without background, 20 films, r=400 μm)

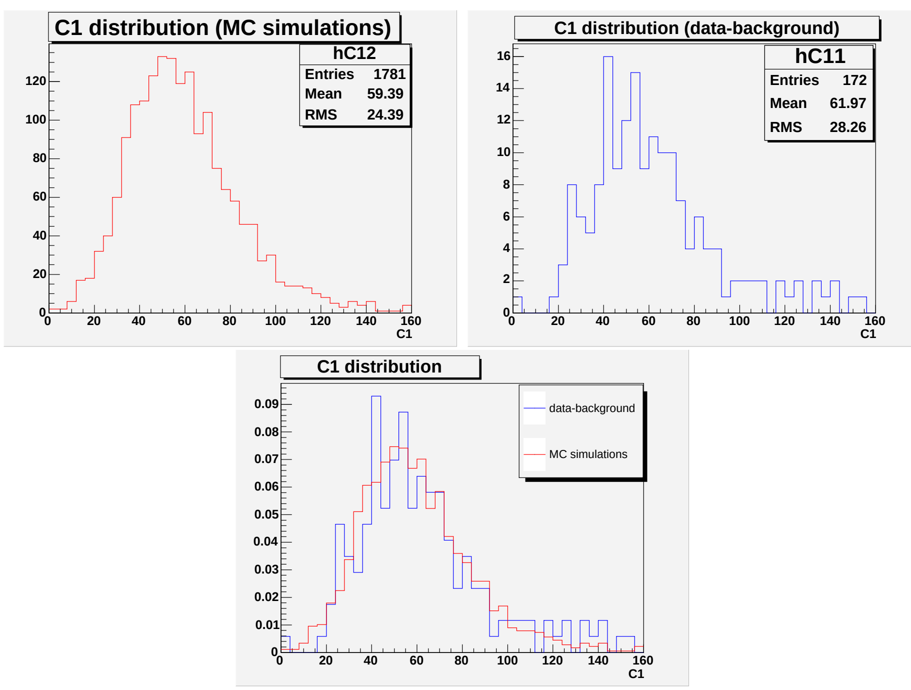

Figure 22: Distributions of the coefficient C1 (20 films, r=400 μm, 6 GeV). The red histogram represents the distribution of simulations withtout background and the blue histogram concerns data after background substraction.

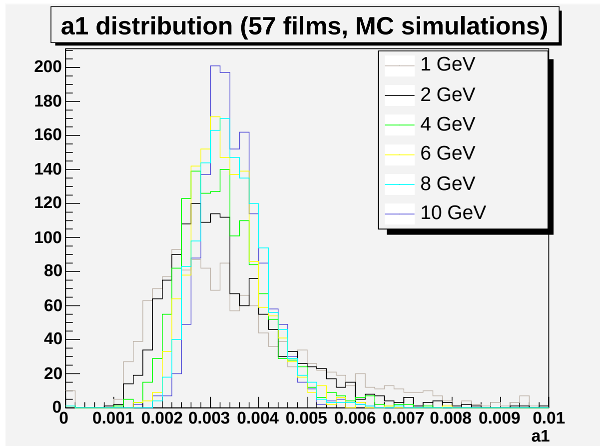

Figure 23: Distributions of the slope a1 (simulations without background, 57 films)

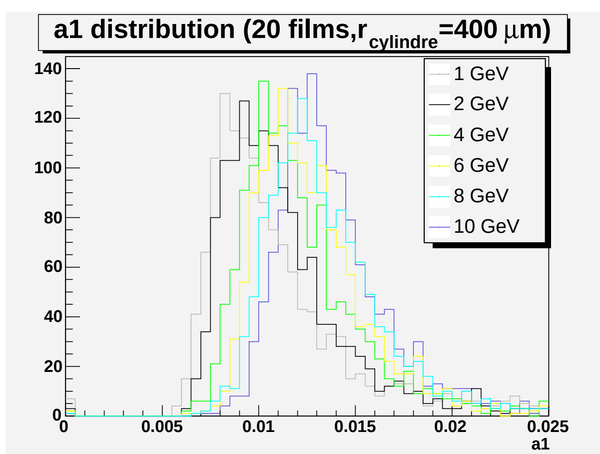

Figure 24: Distribution of the slope a1 (simulations without background, 20 films, r=400 μm)

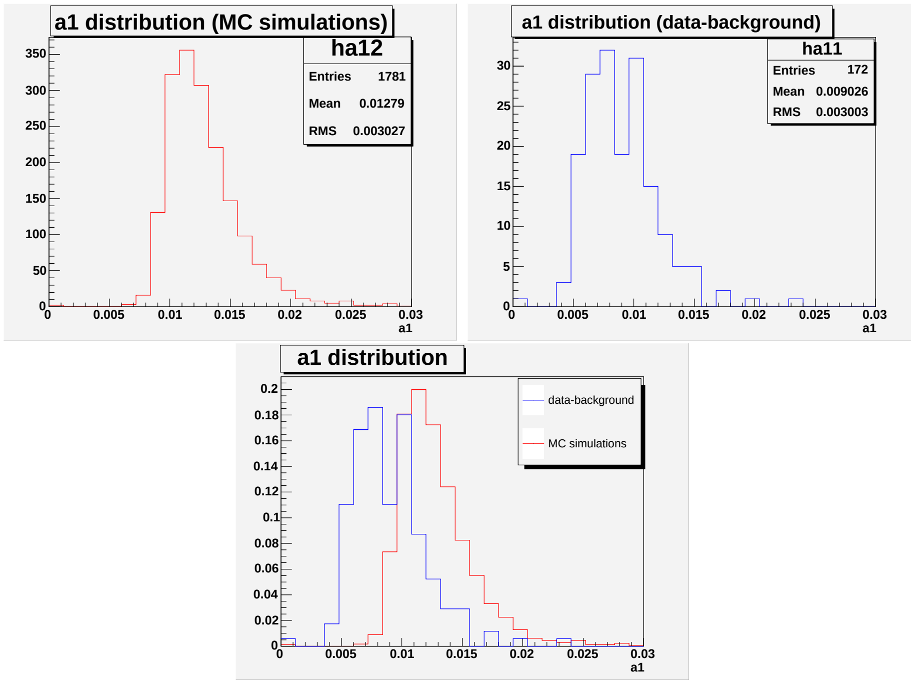

Figure 25: Distributions of the slope a1 (20 films, r=400 μm, 6 GeV). The red histogram represents the distribution of simulations withtout background and the blue histogram concerns data after background substraction.

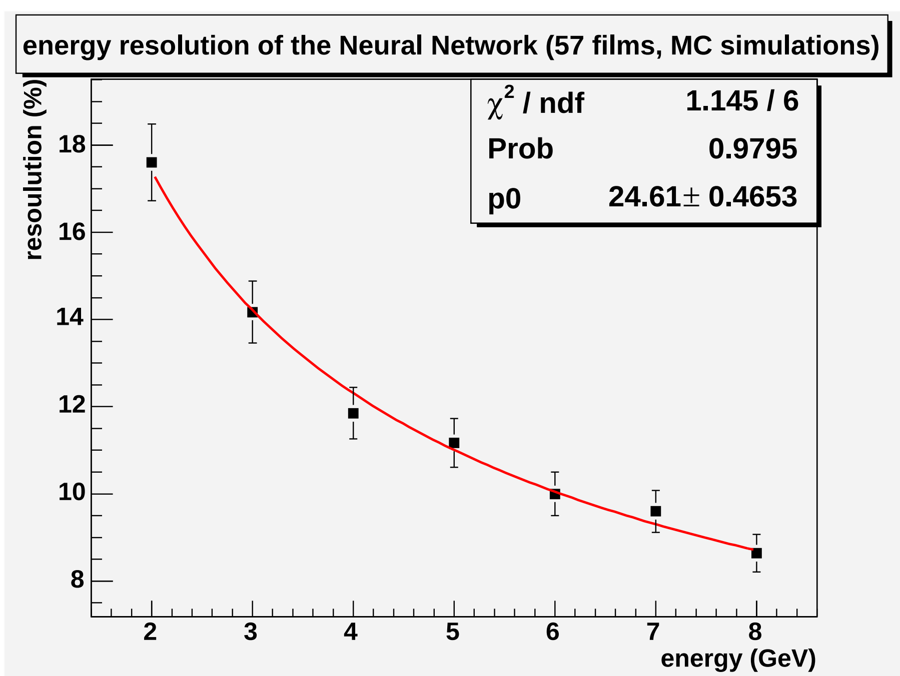

Figure 26: Energy resolution of the Neural Network for electromagnetic cascades developping in 57 emulsions (MC simulations without background)

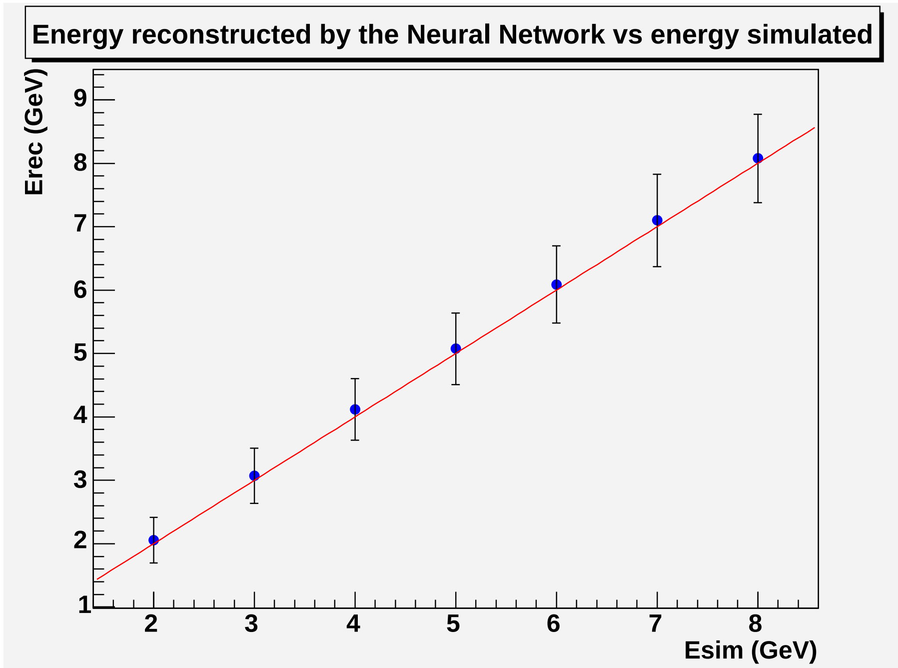

Figure 27: Energy reconstructed versus energy simulated (57 films, MC simulations without background)

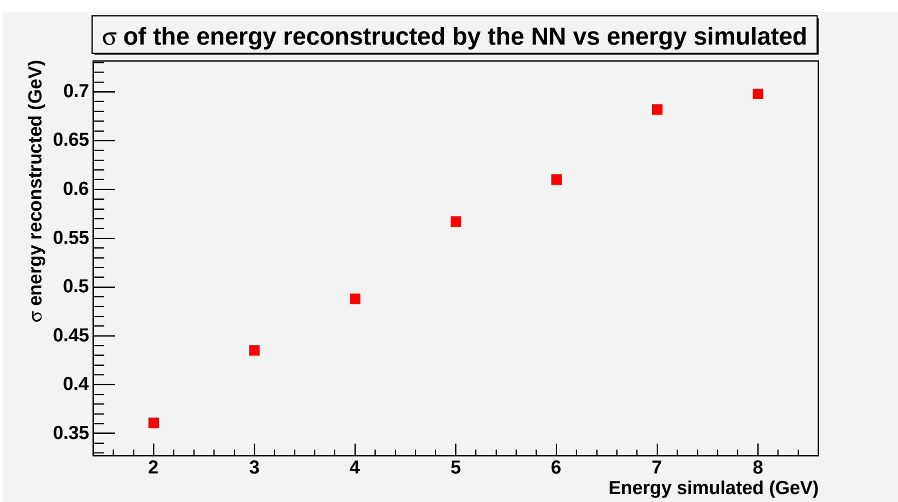

Figure 28: σ on the energy reconstructed by the NN (57 films, MC simulations without background)

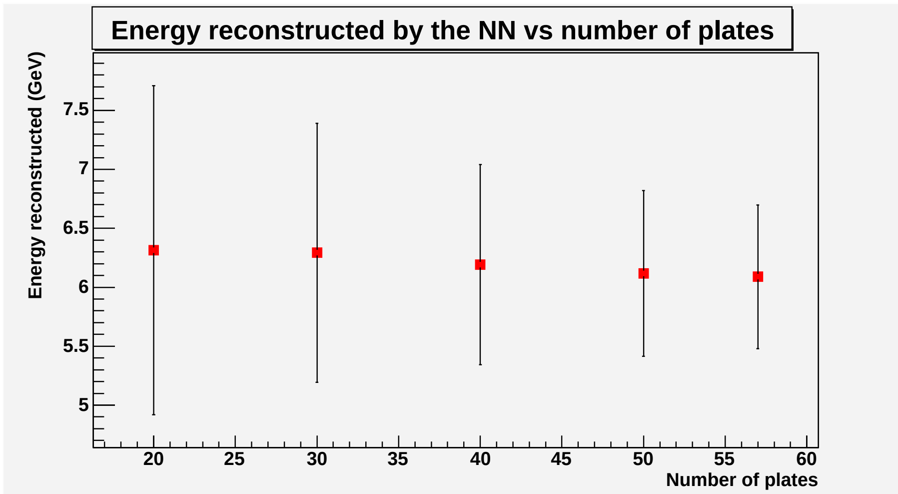

Figure 29: Energy reconstructed by the Neural Network as a function of the number of plates. The true value of the energy simulated is 6 GeV.

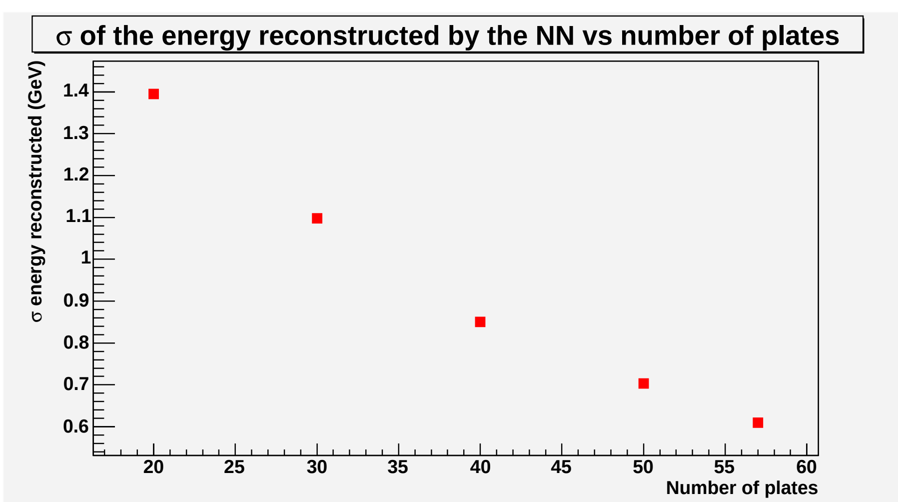

Figure 30: σ on the energy reconstructed by the NN as a function of the number of plates. The true value simulated is 6 GeV (57 films, MC simulations without background)

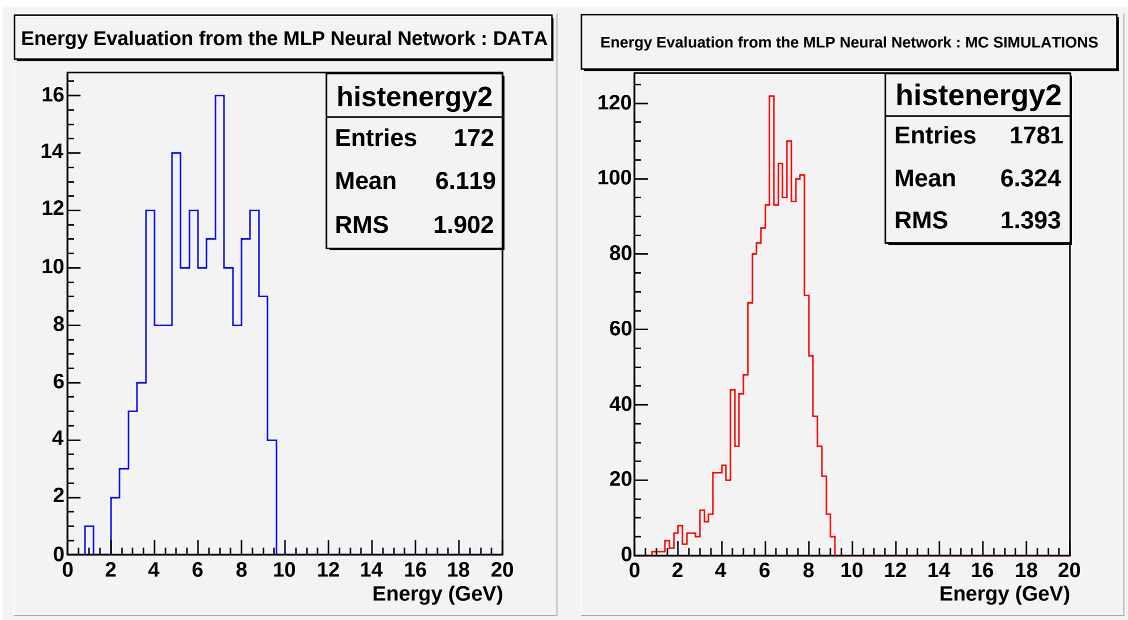

Figure 31: Energy estimation by the Neural Network for experimental data after background substraction (blue) and MC simulations without background (red).